New Trends in Prob. and Stat., Vol. 5, pp. 27–44
T. Kollo *et al.* (Eds)

ON LOCAL DEPENDENCE FUNCTION FOR MULTIVARIATE DISTRIBUTIONS

ISMIHAN BAIRAMOV[1] and SAMUEL KOTZ[2]
[1]Ankara University, Faculty of Science, Department of Statistics, 06100 Tandogan, Ankara, Turkey
e-mail: bayramov@science.edu.tr
[2]Department of Operation Research, George Washington University, Washington, DC
e-mail: kotz@seas.gwu.edu

INTRODUCTION

In recent years several interesting statistical papers have appeared in which scalar association measures are extended to local association functions. Bjerve and Doksum (1993), Doksum *et al.* (1994) and Blyth (1994) introduce and discuss a "correlation curve" which is a generalization of the Pearson correlation coefficient. In particular Bjerve and Doksum (1993) consider a local measure of the strength of the association between random variables Y and X and introduce the correlation curve

$$\rho(x) = \frac{\sigma_1 \beta(x)}{\left[(\sigma_1 \beta(x))^2 + \sigma^2(x)\right]^{1/2}},$$

where $\beta(x) = \mu'(x)$ is the slope of the nonparametric regression $\mu(x) = E(Y \mid X = x)$, $\sigma^2(x) = Var(Y \mid x)$ is the nonparametric residual variance and $\sigma_1^2 = Var(X)$. The idea behind the construction of $\rho(x)$ is based on the fact that in the bivariate normal case $\rho(x) = \rho$ for all x and

$$\rho = \frac{\sigma_1 \beta}{\left[(\sigma_1 \beta)^2 + \sigma^2\right]^{1/2}},$$

where β is the regression slope when Y is regressed on X. The correlation curve $\rho(x)$ is meaningful only when X is a continuous random variable. Jones (1996) provides a motivation for a local dependence function, defined as the mixed partial derivative of the log density, originally proposed by Holland and Wang (1987). There are indeed, many ways of measuring dependence between two random variables. In a recent book Nelsen (1998) widely discusses various measure of dependencies viewing "correlation coefficient" as a measure of the linear dependence between random variables and using the term a "measure of association" for measures such as Kendall's tau and Spearman's rho. Various measures of concordance and their properties are also described in Nelsen's book providing relationship between measures of association and dependence of random variables X and Y.

This work was motivated by the paper of M. C. Jones (1996), (1998) and furnishes motivation for a new local dependence function based on regression concepts. For important bivariate distributions such as bivariate normal, the classical FGM and Sarmanov–Lee class of distributions and several bivariate exponential the expected value of this local dependence function is approximately equal to the Pearson correlation coefficient.

1. THE LOCAL DEPENDENCE FUNCTION

Let (X, Y) be a continuous bivariate random variable with joint cumulative distribution function (d.f.) $F(x, y)$ and with joint probability density function (p.d.f.) $f(x, y)$. The marginal d.f. and p.d.f. of X and Y will as usual be denoted by $F_X(x)$, $f_X(x)$ and $F_Y(x)$, $f_Y(x)$ respectively. Consider the following function of two variables x and y

$$H(x, y) = \frac{E\,(X - E(X \mid Y = y))\,(Y - E(Y \mid X = x))}{\sqrt{E\,(X - E(X \mid Y = y))^2}\sqrt{E\,(Y - E(Y \mid X = x))^2}}, \tag{1.1}$$

which obtained from the expression of the Pearson correlation coefficient by replacing mathematical expectations EX and EY by the conditional expectations $E(X \mid Y = y)$, $E(Y \mid X = x)$ respectively. By construction, $H(x, y)$ can be interpreted as a local dependence function characterizing the dependence between X and Y at a point (x, y). In other words, $H(x, y)$ can characterize the effect (influence) of X on Y "conditionally on X and Y being in a neighborhood of the point (x, y)" and vice versa. After some simple algebraic manipulations (1.1) can be rewritten as follows:

$$H(x, y) = \frac{Cov(X, Y) + \xi_X(y)\xi_Y(x)}{\sqrt{Var(X) + \xi_X^2(y)}\sqrt{Var(Y) + \xi_Y^2(x)}}, \tag{1.2}$$

where $\xi_X(y) = EX - E(X \mid Y = y)$, $\xi_Y(x) = EY - E(Y \mid X = x)$. Dividing the numerator and the denominator of (1.2) by $\sigma_X = \sqrt{Var(X)}$ and $\sigma_Y = \sqrt{Var(Y)}$ one obtains the following expression for $H(x, y)$:

$$H(x, y) = \frac{\rho + \varphi_X(y)\varphi_Y(x)}{\sqrt{1 + \varphi_X^2(y)}\sqrt{1 + \varphi_Y^2(x)}}, \tag{1.3}$$

where $\rho = \frac{Cov(X,Y)}{\sigma_X\sigma_Y}$ (the Pearson correlation coefficient), $\varphi_X(y) = \frac{\xi_X(y)}{\sigma_X}$, $\varphi_Y(x) = \frac{\xi_Y(x)}{\sigma_Y}$. The function $H(x, y)$ will be referred to as a *local dependence function*.

For further analysis of equation (1.3) let $a \leqslant \varphi_X(y) \leqslant b$ and $a \leqslant \varphi_Y(x) \leqslant b$ (possible including $a = -\infty$ and $b = \infty$) for all $(x, y) \in N_{X,Y}$, where $N_{X,Y}$ denotes the support of (X, Y). Consider now the function

$$h(t, z) = \frac{\rho + tz}{\sqrt{1 + t^2}\sqrt{1 + z^2}}, \quad a \leqslant t, z \leqslant b.$$

It is easy to check that

$$h_t = \frac{\partial h}{\partial t} = \frac{z - t\rho}{(1+t^2)^{3/2}\sqrt{1+z^2}}, \qquad h_z = \frac{\partial h}{\partial z} = \frac{t - z\rho}{(1+z^2)^{3/2}\sqrt{1+t^2}}.$$

The critical point of $h(t, z)$, i.e. the solution of $h_t = h_z = 0$, is $t = 0$, $z = 0$. Performing the second derivative test we have:

$$h_{tt} = \frac{-\rho + 2\rho t^2 - 3tz}{(1+t^2)^{5/2}\sqrt{1+z^2}}, \qquad h_{zz} = \frac{-\rho + 2\rho z^2 - 3tz}{(1+z^2)^{5/2}\sqrt{1+t^2}} \quad \text{and}$$

$$h_{tz} = h_{zt} = \frac{1 + \rho tz}{(1+t^2)^{3/2}(1+z^2)^{3/2}}.$$

Consequently at the critical point $(0, 0)$

$$h_{tt}h_{zz} - h_{tz}^2 = \rho^2 - 1 < 0, \quad \text{if } |\rho| < 1.$$

Therefore $h(t, z)$ has a saddle point at $(0, 0)$. That is the point (x^*, y^*) satisfying $\varphi_X(y^*) = \varphi_Y(x^*) = 0$ is a saddle point of $H(x, y)$ and at this point $H(x^*, y^*) = \rho$. It is easy to see that $h(t, z)$ has maximum at the boundary points (a, a) and (b, b) and has minimum at (a, b) and (b, a).

From the equations $\varphi_X(y^*) = \varphi_Y(x^*) = 0$ one obtains $E(X) = E(X \mid Y = y^*)$ and $E(Y) = E(Y \mid X = x^*)$. (Therefore the point (x^*, y^*) satisfying $E(X) = E(X \mid Y = y^*)$ and $E(Y) = E(Y \mid X = x^*)$ is a saddle point of $H(x, y)$ and at this point $H(x, y)$ is equal to ρ.)

From the equations $\varphi_X(y) = a$ and $\varphi_Y(x) = a$ one has $E(X \mid Y = y) = E(X) + a\sigma_X$ and $E(Y \mid X = x) = E(Y) + a\sigma_Y$. Moreover, from the equations $\varphi_X(y) = a$ and $\varphi_Y(x) = b$ one has $E(X \mid Y = y) = E(X) + a\sigma_X$ and $E(Y \mid X = x) = E(Y) + b\sigma_Y$. That is at the point of the maximum of $H(x, y)$ influence of X on Y (and influence of Y on X) is maximal, and conversely at the point of the minimum of $H(x, y)$ the influence of X on Y (and the influence of Y to X) is minimal, as expected.

To simplify calculations below we denote $A_X(y) = E(X \mid Y = y)$ and $A_Y(x) = E(Y \mid X = x)$.

The properties of $H(x, y)$ are given in the following lemma:

LEMMA 1. *The local dependence function $H(x, y)$ has the following properties:*

1°. *If X and Y are independent then $H(x, y) = 0$ for any $(x, y) \in N_{X,Y}$.*

2°. $|H(x, y)| \leqslant 1$, *for all $(x, y) \in N_{X,Y}$.*

3°. *If $|H(x, y)| = 1$ for some $(x, y) \in N_{X,Y}$ then $\rho \neq 0$.*

4°. *Let $A_X(y)$ and $A_Y(x)$ are differentiable functions of their arguments. If $H(x, y) = 0$ for all $(x, y) \in N_{X,Y}$ then $E(X \mid Y = y)$ or $E(Y \mid X = x)$ or both are constant.*

5°. *Let $|\rho| = 1$ and assume that $|H(x, y)| = 1$ at a point (x, y) then $\varphi_X(y) = \varphi_Y(x)$ up to a sign. That is equality of the two standardized distances between conditional and unconditional expectations at (x, y) is an indication that $H(x, y) = 1$ at this point.*

Proof. The property 1° follows directly from (1.1). The property 2° can be obtained from (1.1) using Schwarz inequality.

For proving 3° let $|H(x, y)| = 1$ for some $(x, y) \in N_{X,Y}$. Then from (1.3) one obtains

$$|\rho + \varphi_X(y)\varphi_Y(x)| = \sqrt{1 + \varphi_X^2(y)}\sqrt{1 + \varphi_Y^2(x)}$$

and

$$\rho^2 + 2\rho\varphi_X(y)\varphi_Y(x) = 1 + \varphi_X^2(y) + \varphi_Y^2(x) \tag{1.4}$$

If $\rho = 0$ then from (1.4) one has $\varphi_X^2(y) + \varphi_Y^2(x) = -1$, a contradiction.

For proving 4° let $H(x, y) = 0$ for all $(x, y) \in N_{X,Y}$. Then from (1.3) recalling the definitions $\varphi_X(y)$ and $\varphi_Y(x)$ one can write

$$\rho\sigma_X\sigma_Y = A_Y(x)EX - EXEY - A_X(y)A_Y(x) + EYA_X(y) \tag{1.5}$$

for all $(x, y) \in N_{X,Y}$.

Differentiating (1.5) with respect to x one has

$$(EX - A_X(y))A'_Y(x) = 0 \quad \text{for all } (x, y) \in N_{X,Y}. \tag{1.6}$$

Differentiating (1.6) with respect to y one obtains

$$A'_Y(x)A'_X(y) = 0 \quad \text{for all } (x, y) \in N_{X,Y}.$$

Hence $A_Y(x)$ or $A_X(y)$ or both $A_Y(x)$ and $A_X(y)$ are constant.

For proving 5° let $\rho = 1$ and $|H(x, y)| = 1$. Then from (1.4) one has

$$1 + 2\varphi_X(y)\varphi_Y(x) = 1 + \varphi_X^2(y) + \varphi_Y^2(x)$$

and $(\varphi_X(y) - \varphi_Y(x))^2 = 0$, i.e. $\varphi_X(y) = \varphi_Y(x)$. If $\rho = -1$ then from (1.4) it follows that $(\varphi_X(y) + \varphi_Y(x))^2 = 0$, i.e. $\varphi_X(y) = -\varphi_Y(x)$.

2. EXAMPLES

1. Consider a bivariate normal distribution with joint p.d.f.

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_X^2} - 2\rho\frac{xy}{\sigma_X\sigma_Y} + \frac{y^2}{\sigma_Y^2}\right)\right\},$$

where $\sigma_X^2 = Var(X)$, $\sigma_Y^2 = Var(Y)$, $EX = 0$, $EY = 0$, $\rho = \frac{Cov(X,Y)}{\sigma_X\sigma_Y}$. It is evident that $E(Y \mid X = x) = \rho\frac{\sigma_Y}{\sigma_X}x$, $E(X \mid Y = y) = \rho\frac{\sigma_X}{\sigma_Y}y$ and

$$H(x, y) = \frac{\rho\sigma_X\sigma_Y + \rho^2 xy}{\sqrt{\sigma_X^2 + \rho^2x^2}\sqrt{\sigma_Y^2 + \rho^2y^2}}.$$

In the standardized case when $\sigma_X = \sigma_Y = 1$ we have

$$H(x, y) = \frac{\rho + \rho^2 xy}{\sqrt{1 + \rho^2 y^2}\sqrt{1 + \rho^2 x^2}}. \tag{2.1}$$

Consider (2.1). It is clear that $H(0, 0) = \rho$, i.e. the Pearson correlation coefficient corresponds to the local dependence at the point $(x, y) = (0, 0)$. We provide some values of $H(x, y)$ for selected x and y.

x	y	$H(x, y)$	x	y	$H(x, y)$
0	0	ρ	1	1	$\frac{\rho + \rho^2}{1 + \rho^2}$
0	1	$\frac{\rho}{\sqrt{1 + \rho^2}}$	10	10	$\frac{\rho + 100\rho^2}{1 + 100\rho^2}$
0	2	$\frac{\rho}{\sqrt{1 + 4\rho^2}}$	−1	0	$\frac{\rho}{\sqrt{1 + \rho^2}}$
0	10	$\frac{\rho}{\sqrt{1 + 100\rho^2}}$	−1	1	$\frac{\rho - \rho^2}{1 + \rho^2}$
−3	3	$\frac{\rho - 9\rho^2}{1 + 9\rho^2}$	−2	−2	$\frac{\rho + 4\rho^2}{1 + 4\rho^2}$
−3	1	$\frac{\rho - 3\rho^2}{\sqrt{1 + 9\rho^2}\sqrt{1 + \rho^2}}$	−3	−3	$\frac{\rho + 9\rho^2}{1 + 9\rho^2}$
−5	0	$\frac{\rho}{\sqrt{1 + 25\rho^2}}$	−4	−4	$\frac{\rho + 16\rho^2}{1 + 16\rho^2}$
−10	10	$\frac{\rho - 100\rho^2}{1 + 100\rho^2}$	−5	−5	$\frac{\rho + 25\rho^2}{1 + 25\rho^2}$

One observes from the table above that $H(x, y)$ takes large values when (x, y) lies on the diagonal $x = y$ and can be seen from the graph that $x = y$ implies $H(x, y) \uparrow$. Conversely if $x = -y$ and $|x| \uparrow$ then $H(x, y) \downarrow$. Let $\rho = 0.99$. That is X and Y are nearly linearly dependent. In the Table 1 some numerical values of $H(x, y)$ for selected x and y for this ρ are given.

Note that the rate of decrease of $H(x, y)$ for $x = 0$ as y deviates from 0 becomes smaller for larger y. For $x \neq 0$ its deviation from y does not affect significantly the value of $H(x, y)$ (which of course decreases as discrepancy increases).

Table 1.
Numerical values of $H(x, y)$ for $\rho = 0.99$

x	y	$H(x, y)$	x	y	$H(x, y)$	x	y	$H(, y)$	x	y	$H(x, y)$
0	0	0.99	0	9	0.11	1	100	0.71	−10	10	−0.98
0	1	0.70	0	10	0.09	2	2	1.00	−1	2	−0.31
0	2	0.45	0	20	0.05	2	10	0.93	−10	10	−0.98
0	3	0.32	0	50	0.02	5	10	0.99	−15	0	0.07
0	4	0.24	0	100	0.01	9	10	1.00	−15	1	−0.65
0	5	0.20	0	1000	0.00	10	10	1.00	−15	5	−0.96
0	6	0.16	1	1	0.99	−1	0	0.70	−15	15	−0.99
0	7	0.14	1	2	0.95	−10	0	0.10	−15	−15	1.00
0	8	0.12	1	10	0.77	−1	1	0.01	−20	20	−0.99

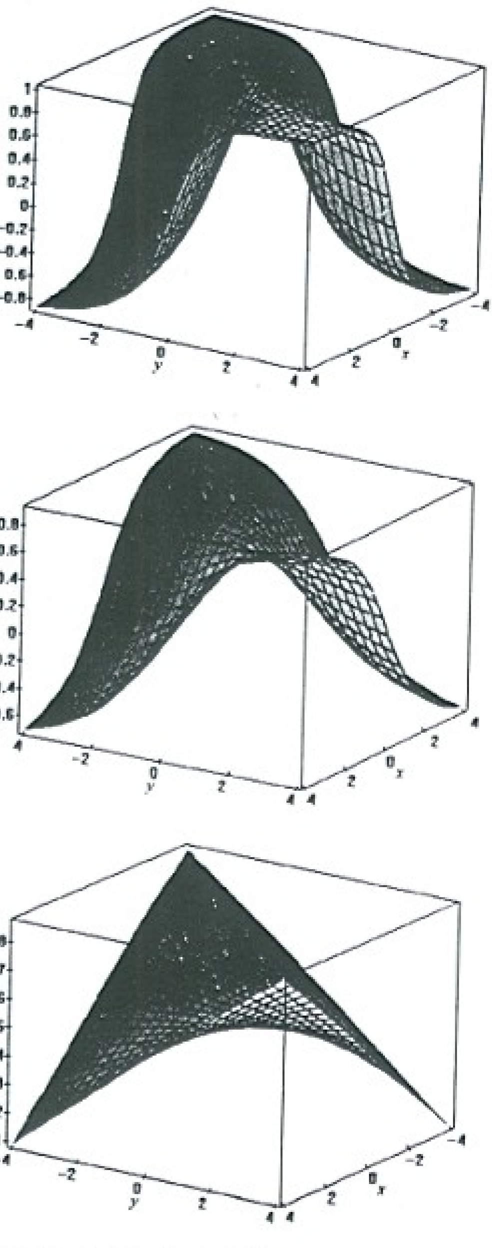

Figure 1. Graph of $H(x, y)$ given in (2.1) for $\rho = 0.95$; 0.5; 0.05 versus $-4 \leqslant x, y \leqslant 4$.

Table 2.
Numerical values of $H(x, y)$ for $\rho = 0.1$

x	y	$H(x, y)$	x	y	$H(x, y)$
0	0	0.10	10	10	0.55
0	1	0.09	10	100	0.71
0	10	0.07	10	1000	0.71
0	100	0.01	100	100	0.99
1	1	0.11	−10	−10	0.55
1	10	0.14	−10	10	−0.45
1	20	0.13	−20	0	0.04
1	50	0.12	−20	20	−0.78
1	100	0.11	−100	100	−0.99

Now let $\rho = 0.1$. The numerical values of $H(x, y)$ are given in the Table 2.

Comparing two tables above for $\rho = 0.9$ and $\rho = 0.1$ one observes that large values of ρ dependence at any point (x, y) is stronger than for small values of ρ. More precisely the following can be asserted. For $x = y = 0$ the values of $H(x, y) \leqslant \rho$ and slowly monotonically decrease as $x = 0$ but $y \uparrow$. Already for $x = y = 1$, $H(x, y) > \rho$ and for $x = 1$ and increasing y, $H(x, y)$ is not monotonic. However $H(x, y) > \rho$ is valid. For large $x = y = 10$ $H(x, y)$ is substantially larger than ρ and increases as y deviates from the value $x = 10$. Note that for $\rho = 0.1$ while $H(x, y) = 0.991$ for $x = y = 100$. This intricate relationship between ρ and $H(x, y)$ further justifies the importance of a measure of local dependence. The rather complex relation between ρ and $H(x, y)$ for x and y of given different signs can be analyzed from the numerical values in Table 1 and Table 2.

Computation of $EH(X, Y)$ is rather cumbersome in this case. We provide some numerical values for the integral

$$J(\rho) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \int_{-4}^{4}\int_{-4}^{4} \frac{\rho + \rho^2 xy}{\sqrt{1+\rho^2y^2}\sqrt{1+\rho^2x^2}} \times \exp\left\{-\frac{1}{2(1-\rho^2)}\left(x^2 - 2\rho xy + y^2\right)\right\} \mathrm{d}x\mathrm{d}y$$

computed with the aid of "MATHCAD":

ρ	±0.100	±0.200	±0.300	±0.400	±0.500	±0.600	±0.700	±0.800	±0.900
$J(\rho)$	±0.100	±0.200	±0.299	±0.398	±0.495	±0.592	±0.689	±0.788	±0.890

Indeed there is no discrepancy (whatsoever at least up to 0.3 significant figures) between ρ and $J(\rho)$ for values of ρ less than 3 and even for ρ as high as 0.9 and $J(\rho) = 0.890$. We note however that the relation $J(\rho) \leqslant \rho$ holds throughout. The same result is valid when $J(\rho)$ was calculated with the range of intervals taken from −10 to 10.

2. Consider now the one parameter family of FGM (Farlie–Gumbel–Morgenstern) distributions with uniform marginals. The corresponding d.f. is

$$F_\alpha(x, y) = xy\{1 + \alpha(1 - x)(1 - y)\};$$

and the joint p.d.f. is given by

$$f_\alpha(x, y) = 1 + \alpha(1 - 2x)(1 - 2y), \quad 0 \leqslant x, y \leqslant 1, \ -1 \leqslant \alpha \leqslant 1.$$

For generalizations and further discussion of this family see e.g. Johnson and Kotz (1975), (1977).

It is evident that $E(Y \mid X = x) = \frac{1}{2} - \frac{1}{6}\alpha(1 - 2x)$, $E(X \mid Y = y) = \frac{1}{2} - \frac{1}{6}\alpha(1 - 2y)$, $EX = EY = \frac{1}{2}$, $Var(X) = Var(Y) = \frac{1}{12}$, $Cov(X, Y) = EXY - EXEY = \frac{1}{36}\alpha$ and $\rho = \frac{\alpha}{3}$. Thus,

$$H(x, y) = \frac{\alpha + \alpha^2(1 - 2x)(1 - 2y)}{\sqrt{3 + \alpha^2(1 - 2x)^2}\sqrt{3 + \alpha^2(1 - 2y)^2}}. \tag{2.2}$$

Consider (2.2). One observes that in the point of symmetry $(x, y) = (\frac{1}{2}, \frac{1}{2})$ the local dependence function is as before equal to the correlation coefficient, i.e.

$$H\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{\alpha}{3} = \rho.$$

We now provide some values of $H(x, y)$ for fixed (x, y).

x	y	$H(x, y)$	x	y	$H(x, y)$
0	0	$\frac{\alpha + \alpha^2}{3 + \alpha^2}$	1	1	$\frac{\alpha + \alpha^2}{3 + \alpha^2}$
0	1	$\frac{\alpha - \alpha^2}{3 + \alpha^2}$	1	0	$\frac{\alpha - \alpha^2}{3 + \alpha^2}$
0	$\frac{1}{2}$	$\frac{\alpha}{\sqrt{3 + \alpha^2}\sqrt{3}}$	$\frac{1}{2}$	0	$\frac{\alpha}{\sqrt{3 + \alpha^2}\sqrt{3}}$
0	$\frac{1}{3}$	$\frac{3\alpha + \alpha^2}{\sqrt{3 + \alpha^2}\sqrt{27 + \alpha^2}}$	$\frac{1}{3}$	0	$\frac{3\alpha + \alpha^2}{\sqrt{3 + \alpha^2}\sqrt{27 + \alpha^2}}$

Maximal $H(x, y)$ is attained at $(x, y) = (0, 0)$ or $(x, y) = (1, 1)$. Here $H(x, y) > \frac{\alpha}{3}$ with a dip at the saddle point $(x, y) = \frac{1}{2}$, here $H(x, y) = \frac{\alpha}{3}$. The values $H(x, y)$ for the asymmetric points such as $(0, 1)$, $(1, 0)$ are the same and smaller than $\frac{\alpha}{3}$.

The expected value of $H(x, y)$ is relatively straightforward

$$E[H(X, Y)] \equiv E(\alpha) = \int_0^1 \int_0^1 H(x, y) f_\alpha(x, y)\, dx\, dy.$$

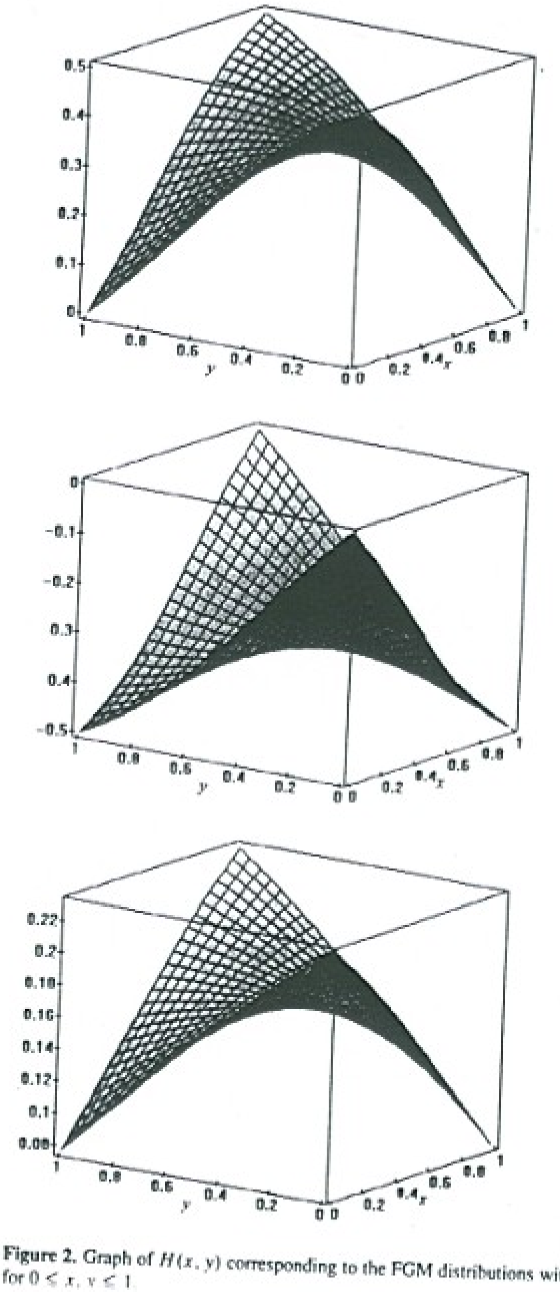

Figure 2. Graph of $H(x, y)$ corresponding to the FGM distributions with $\alpha = 1; -1; 0.5$ for $0 \leq x, y \leq 1$.

Using (2.2) and the definition of $f_\alpha(x, y)$ one can write

$$E(\alpha)$$

$$= \int_0^1\int_0^1 \frac{\alpha + \alpha^2(1-2x)(1-2y)}{\sqrt{3+\alpha^2(1-2x)^2}\sqrt{3+\alpha^2(1-2y)^2}}\,[1+\alpha(1-2x)(1-2y)]\,\mathrm{d}x\,\mathrm{d}y$$

$$= \int_0^1\int_0^1 \frac{\alpha + \alpha^2(1-2x)(1-2y) + \alpha^2(1-2x)(1-2y) + \alpha^3(1-2x)(1-2y)}{\sqrt{3+\alpha^2(1-2x)^2}\sqrt{3+\alpha^2(1-2y)^2}}\,\mathrm{d}x\,\mathrm{d}y$$

$$= \alpha\int_0^1\int_0^1 \frac{\mathrm{d}x\,\mathrm{d}y}{\sqrt{3+\alpha^2(1-2x)^2}\sqrt{3+\alpha^2(1-2y)^2}}$$

$$+ 2\alpha^2\int_0^1\int_0^1 \frac{(1-2x)(1-2y)\,\mathrm{d}x\,\mathrm{d}y}{\sqrt{3+\alpha^2(1-2x)^2}\sqrt{3+\alpha^2(1-2y)^2}}$$

$$+ \alpha^3\int_0^1\int_0^1 \frac{(1-2x)^2(1-2y)^2\,\mathrm{d}x\,\mathrm{d}y}{\sqrt{3+\alpha^2(1-2x)^2}\sqrt{3+\alpha^2(1-2y)^2}}. \tag{2.3}$$

One can check that

$$\int_0^1 \frac{\mathrm{d}x}{\sqrt{3+\alpha^2(1-2x)^2}} = \frac{1}{2\alpha}\ln\left(\frac{\sqrt{3+\alpha^2}+\alpha}{\sqrt{3+\alpha^2}-\alpha}\right), \tag{2.4}$$

$$\int_0^1 \frac{(1-2x)}{\sqrt{3+\alpha^2(1-2x)^2}}\,\mathrm{d}x = 0, \tag{2.5}$$

$$\int_0^1 \frac{(1-2x)^2}{\sqrt{3+\alpha^2(1-2x)^2}}\,\mathrm{d}x = \frac{1}{4\alpha^3}\left[2\alpha\sqrt{3+\alpha^2} - 3\ln\left(\frac{\sqrt{3+\alpha^2}+\alpha}{\sqrt{3+\alpha^2}-\alpha}\right)\right]. \tag{2.6}$$

Using (2.4), (2.5), (2.6) in (2.3) one obtains

$$E(\alpha) = \frac{1}{4\alpha}\left[\ln\left(\frac{\sqrt{3+\alpha^2}+\alpha}{\sqrt{3+\alpha^2}-\alpha}\right)\right]^2 + \frac{1}{16\alpha^3}\left[2\alpha\sqrt{3+\alpha^2} - 3\ln\left(\frac{\sqrt{3+\alpha^2}+\alpha}{\sqrt{3+\alpha^2}-\alpha}\right)\right]^2. \tag{2.7}$$

The Taylor expansion of $E[H(X, Y)]$ around $\alpha = 0$ is

$$E(\alpha) = \frac{\alpha}{3} - \frac{\alpha^5}{1215} + \frac{4}{14175}\alpha^7 - \frac{32}{382725}\alpha^9 + o(\alpha^{11}). \tag{2.8}$$

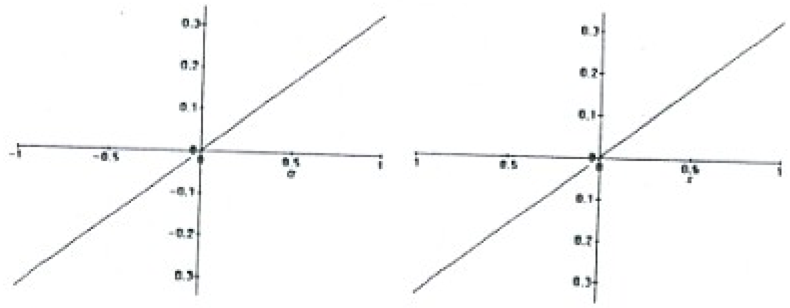

Figure 3. Graph of $E(\alpha)$ versus $-1 \leqslant \alpha \leqslant 1$ and graph of $f(x) = \frac{x}{3}$, $-1 \leqslant x \leqslant 1$ (for comparison).

Note that $E(\alpha) = \frac{\alpha}{3} = \rho$ up to the order α^4.

Figure 3 compares the graphs of $E(\alpha)$ over $-1 \leqslant \alpha \leqslant 1$ and $f(\alpha) = \frac{1}{3}\alpha$ over the same range of values of α.

We provide some numerical values for $E(\alpha)$:

α	$E(\alpha)$	α	$E(\alpha)$
± 1	± 0.333	± 0.5	± 0.167
± 0.9	± 0.3	± 0.4	± 0.133
± 0.8	± 0.266	± 0.3	± 0.1
± 0.7	± 0.233	± 0.2	± 0.067
± 0.6	± 0.2	± 0.1	± 0.033

Therefore $E(\alpha) = E[H(X, Y)] = \frac{\alpha}{3} = \rho$ up to the third significant figure. We are unable yet to assert that $E(\alpha) \leqslant \frac{\alpha}{3}$. Now we provide some numerical values for $H(x, y)$ and for comparison the joint p.d.f. $f_\alpha(x, y)$ for $\alpha = 0.99$, see Table 3.

One can observe from the table above that small values of $f_\alpha(x, y)$ correspond to the small values of $H(x, y)$ and vice versa. Roughly speaking, $f_\alpha(x, y)$ is the probability that (X, Y) takes values at the neighborhood of (x, y) and the conditional probability of the event "X takes values at the neighborhood of x, with the condition that Y takes values at the neighborhood of y" is expressed by $f_\alpha(x, y)$. That is, small value of $f_\alpha(x, y)$ corresponds to a weak dependence of X and Y at the point (x, y).

Table 3.
Some numerical values of $H(x, y)$ and p.d.f. $f(x, y)$ for $\alpha = 0.99$

ρ	α	x	y	$f(x, y)$	$H(x, y)$	α	x	y	$f(x, y)$	$H(x, y)$
0.33	0.99	0.1	0.1	1.63	0.44	0.99	0.2	0.4	1.12	0.29
–	–	0.1	0.2	1.42	0.41	–	0.4	0.6	0.96	0.31
–	–	0.1	0.3	1.32	0.39	–	0.5	0.5	1	0.33
–	–	0.1	0.4	1.15	0.35	–	0.6	0.7	1.10	0.34
–	–	0.1	0.6	0.84	0.25	–	0.6	0.9	1.15	0.35
–	–	0.1	0.9	0.37	0.1	–	0.9	0.9	1.63	0.45

Conversely, large value of $f_\alpha(x, y)$ corresponds to a strong dependence at the point (x, y). Observe also that the values of f and H are sensitive to the distances between the corresponding x and y. Comparisons given in Table 3 provide yet another motivation of using above $H(x, y)$ as a local dependence function.

3. Consider the Sarmanov–Lee bivariate density introduced by Lee (1996),

$$f_\alpha(x, y) = f_X(x) f_Y(y) \{1 + \alpha \psi_1(x) \psi_2(y)\}, \tag{2.9}$$

where $\psi_1(x) = x - \mu_X$, $\psi_2(y) = y - \mu_Y$, $\mu_X = EX$, $\mu_Y = EY$. For uniform marginals, (2.9) has the form

$$f_\alpha(x, y) = 1 + \alpha\left(x - \frac{1}{2}\right)\left(y - \frac{1}{2}\right), \quad -4 \leqslant \alpha \leqslant 4. \tag{2.10}$$

(Compare with the density corresponding to the p.d.f. of the FGM distribution with uniform marginals given in the previous section). It is easy to observe that for $f_\alpha(x, y)$ in (2.10) $E(X \mid Y = y) = \frac{1}{2} + \frac{\alpha}{12}(y - \frac{1}{2})$, $E(Y \mid X = x) = \frac{1}{2} + \frac{\alpha}{12}(x - \frac{1}{2})$, $\rho = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\alpha}{12}$. Thus

$$H(x, y) = \frac{\alpha + \alpha^2(x - \frac{1}{2})(y - \frac{1}{2})}{\sqrt{12 + \alpha^2(x - \frac{1}{2})^2}\sqrt{12 + \alpha^2(y - \frac{1}{2})^2}}.$$

Consider

$$\begin{aligned} E[H(X, Y)] \equiv E_2(\alpha) &= \int_0^1 \int_0^1 H(x, y) f_\alpha(x, y)\, dx\, dy \\ &= \int_0^1 \int_0^1 \frac{\left[\alpha + \alpha^2(x - \frac{1}{2})(y - \frac{1}{2})\right]\left[1 + \alpha(x - \frac{1}{2})(y - \frac{1}{2})\right]}{\sqrt{12 + \alpha^2(x - \frac{1}{2})^2}\sqrt{12 + \alpha^2(y - \frac{1}{2})^2}}\, dx\, dy. \end{aligned}$$

One has

$$\begin{aligned} E_2(\alpha) &= \int_0^1 \int_0^1 \frac{\alpha}{\sqrt{12 + \alpha^2(x - \frac{1}{2})^2}\sqrt{12 + \alpha^2(y - \frac{1}{2})^2}}\, dx\, dy \\ &\quad + 2\alpha^2 \int_0^1 \int_0^1 \frac{(x - \frac{1}{2})(y - \frac{1}{2})}{\sqrt{12 + \alpha^2(x - \frac{1}{2})^2}\sqrt{12 + \alpha^2(y - \frac{1}{2})^2}}\, dx\, dy \\ &= \alpha^3 \int_0^1 \int_0^1 \frac{(x - \frac{1}{2})^2(y - \frac{1}{2})^2}{\sqrt{12 + \alpha^2(x - \frac{1}{2})^2}\sqrt{12 + \alpha^2(y - \frac{1}{2})^2}}\, dx\, dy. \end{aligned} \tag{2.11}$$

One can check that

$$\int_0^1 \frac{1}{\sqrt{12 + \alpha^2(x - \frac{1}{2})^2}}\, dx = \frac{1}{\alpha^2} \ln \frac{\sqrt{48 + \alpha^2} + \alpha}{\sqrt{48 + \alpha^2} - \alpha}, \tag{2.12}$$

$$\int_0^1 \frac{x-\frac{1}{2}}{\sqrt{12+\alpha^2(x-\frac{1}{2})^2}}\,\mathrm{d}x = 0, \tag{2.13}$$

$$\int_0^1 \frac{(x-\frac{1}{2})^2}{\sqrt{12+\alpha^2(x-\frac{1}{2})^2}}\,\mathrm{d}x = \frac{1}{4\alpha^3}\left[\alpha\sqrt{48+\alpha^2} - 24\ln\frac{\sqrt{48+\alpha^2}+\alpha}{\sqrt{48+\alpha^2}-\alpha}\right]. \tag{2.14}$$

Therefore from (2.10), (2.11), (2.12) and (2.13) one has

$$E_2(\alpha) = \frac{1}{\alpha}\left[\ln\frac{\sqrt{48+\alpha^2}+\alpha}{\sqrt{48+\alpha^2}-\alpha}\right]^2 + \frac{1}{16\alpha^3}\left[\alpha\sqrt{48+\alpha^2} - 24\ln\frac{\sqrt{48+\alpha^2}+\alpha}{\sqrt{48+\alpha^2}-\alpha}\right]^2.$$

The Taylor expansion of $E_2(\alpha)$ about $\alpha = 0$ is

$$E_2(\alpha) = \frac{\alpha}{12} - \frac{\alpha^5}{1\,244\,160} + \frac{\alpha^7}{58\,060\,800} + o(\alpha^9). \tag{2.15}$$

Therefore $E_2(\alpha)$ is approximately equal to $\rho = \frac{\alpha}{12}$. In fact comparing (2.8) and (2.15) we observe that for the Sarmanov family $E[H(X,Y)]$ is even closer to the value of ρ than those for the FGM uniform distribution.

We conclude by providing two examples dealing with bivariate distributions with exponential marginal.

4. Consider the following bivariate distribution of exponential type due to Abrahams and Thomas (1994) and Arnold and Strauss (1988) with joint p.d.f.

$$f_A(x,y) = k\exp(-x-y-\theta xy), \quad 0 < x, y < \infty,\ \theta \geqslant 0, \tag{2.16}$$

where $k = \theta\exp(-\frac{1}{\theta})/\mathrm{Ei}(\frac{1}{\theta})$ and $\mathrm{Ei}(u) := \int_u^\infty v^{-1}\mathrm{e}^{-v}\mathrm{d}v$. An important feature of this distribution is that the conditional distributions are exponential distributions while the marginals are of a simple form proportional to $(1+\theta x)^{-1}\mathrm{e}^{-x}$:

$$f_X(x) = k\mathrm{e}^{-x}(1+\theta x), \qquad f_Y(y) = k\mathrm{e}^{-y}(1+\theta y).$$

Below we present graphs of $f_A(x,y)$ for $\theta = 1$ and $\theta = 10$.

It is evident that

$$f(x \mid Y = y) = \frac{f_A(x,y)}{f_Y(y)} = (1+\theta y)\mathrm{e}^{-x(1+\theta y)}$$

and

$$f(y \mid X = x) = \frac{f_A(x,y)}{f_X(x)} = (1+\theta x)\mathrm{e}^{-y(1+\theta x)}.$$

Therefore

$$E(X \mid Y = y) = \int_0^\infty x(1+\theta y)\mathrm{e}^{-x(1+\theta y)}\,\mathrm{d}x = (1+\theta y)^{-1},$$

$$E(Y \mid X = x) = \int_0^\infty y(1+\theta x)\mathrm{e}^{-y(1+\theta x)}\,\mathrm{d}y = (1+\theta x)^{-1}.$$

One can easily check that

$$EX = k\int_0^\infty x(1+\theta x)^{-1}e^{-x}\,dx = \left(\theta - \mathrm{Ei}\left(\frac{1}{\theta}\right)e^{\frac{1}{\theta}}\right)\frac{k}{\theta^2} = \frac{k-1}{\theta}, \quad EY = \frac{k-1}{\theta}$$

and

$$EX^2 = k\int_0^\infty x^2(1+\theta x)^{-1}e^{-x}\,dx = \left(\mathrm{Ei}\left(\frac{1}{\theta}\right)e^{\frac{1}{\theta}} + \theta^2 - \theta\right)\frac{k}{\theta^3} = k\left(\frac{1}{\theta} - \frac{1}{\theta^2}\right) + \frac{1}{\theta^2}$$

(note the definition of k presented above). Therefore

$$Var(X) = Var(Y) = k\left(\frac{1}{\theta} + \frac{1}{\theta^2}\right) - \frac{k^2}{\theta^2}.$$

Also EXY:

$$\begin{aligned} EXY &= k\int_0^\infty\int_0^\infty xye^{-x-y-\theta xy}\,dx\,dy = k\int_0^\infty\int_0^\infty xye^{-y}e^{-x(1+\theta y)}\,dx\,dy \\ &= k\int_0^\infty ye^{-y}\left[\int_0^\infty xe^{-x(1+\theta y)}\,dx\right]dy \\ &= k\int_0^\infty \frac{ye^{-y}}{(1+\theta y)^2}\,dy = \frac{k}{\theta^3}\left(\theta\mathrm{Ei}\left(\frac{1}{\theta}\right)e^{\frac{1}{\theta}} + \mathrm{Ei}\left(\frac{1}{\theta}\right)e^{\frac{1}{\theta}} - \theta\right) = \frac{1-k}{\theta^2} + \frac{1}{\theta}. \end{aligned}$$

Therefore

$$Cov(X,Y) = EXY - EXEY = \frac{k-k^2}{\theta^2} + \frac{1}{\theta} \quad \text{and}$$

$$\rho = \frac{k-k^2+\theta}{k\theta + k - k^2}.$$

We have

$$\xi_X(y) = EX - E(X \mid Y = y) = \frac{k}{\theta} - \frac{1}{\theta} - \frac{1}{1+\theta y} \quad \text{and}$$

$$\xi_Y(x) = EY - E(Y \mid X = x) = \frac{k}{\theta} - \frac{1}{\theta} - \frac{1}{1+\theta x}.$$

Then using (1.2) the local dependence function can be written as follows:

$$H(x,y) = \frac{\frac{1}{\theta^2}(k-k^2+\theta) + \left(\frac{k-1}{\theta} - \frac{1}{1+\theta y}\right)\left(\frac{k-1}{\theta} - \frac{1}{1+\theta x}\right)}{\sqrt{k\frac{1+\theta-k}{\theta^2} + \left(\frac{k-1}{\theta} - \frac{1}{1+\theta y}\right)^2}\sqrt{k\frac{1+\theta-k}{\theta^2} + \left(\frac{k-1}{\theta} - \frac{1}{1+\theta x}\right)^2}}. \tag{2.17}$$

Calculations show that the $x^* = y^* = 0.853$ is the saddle point of (2.17) for $\theta = 0.1$ and $x^* = y^* = 0.477$ is the corresponding point for $\theta = 1$. Graphical representations of the density (2.16) (not available in the literature) are presented in Figure 4.

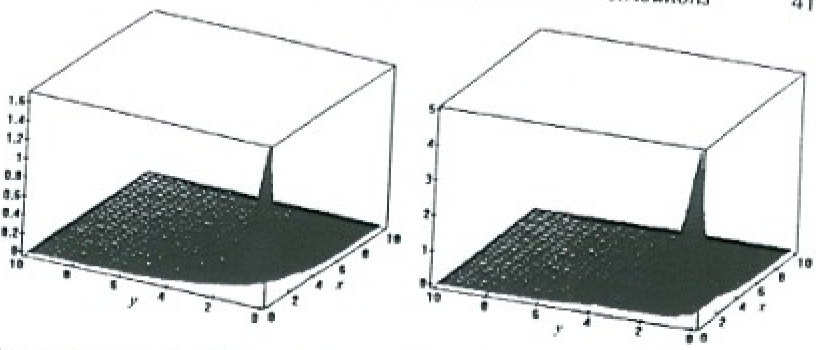

Figure 4. Graph of $f_A(x, y)$ for $\theta = 1$ and $\theta = 10$ versus $1 \leqslant x, y \leqslant 10$.

Table 4.
Some numerical values of $H(x, y)$ for $f_a(x, y)$ in (2.9) for $\theta = 1$

x	y	$H(x, y)$	x	y	$H(x, y)$	x	y	$H(x, y)$	x	y	$H(x, y)$
0	0	−0.047	1	1	−0.181	2	2	−0.026	4	4	0.12
0	1	−0.316	1	2	−0.121	2	4	0.04	4	5	0.138
0	2	−0.377	1	3	−0.092	2	10	0.087	4	10	0.175
0	3	−0.399	1	4	−0.076	2	20	0.103	5	5	0.156
0	4	−0.41	1	5	−0.066	2	100	0.117	5	6	0.169
0	5	−0.417	1	10	−0.044	3	3	0.065	5	10	0.195
0	10	−0.429	1	20	−0.032	3	5	0.092	6	6	0.182
0	100	−0.439	1	100	−0.023	3	10	0.144	6	10	0.208

In the Table 4 are some numerical values for $H(x, y)$:

$$\theta = 1 \ (k = 1,677, \rho = -0.249, x^* = y^* = 0.477, H(x^*, y^*) = -0.249)$$

Note that although ρ is negative, in this case $H(x, y)$ may obtain positive values for large values of x and y. For small values of x and y $H(x, y)$ is negative and its absolute value may exceed the absolute value of ρ.

5. Finally consider Gumbel's bivariate exponential distribution (Gumbel 1960):

$$F_{X,Y}(x, y) = 1 - e^{-x} - e^{-y} + e^{-(x+y+\theta xy)}, \quad x, y > 0, \ 0 \leqslant \theta \leqslant 1.$$

The p.d.f. is

$$f_{X,Y}(x, y) = e^{-x-y-\theta xy}\{(1 + \theta x)(1 + \theta y) - \theta\}.$$

One can check that

$$E(Y \mid X = x) = (1 + \theta + \theta x)(1 + \theta x)^{-2} \quad \text{and}$$

$$E(X \mid Y = y) = (1 + \theta + \theta y)(1 + \theta y)^{-2}.$$

The Pearson correlation coefficient is:

$$\rho = \theta^{-1} e^{\frac{1}{\theta}} \operatorname{Ei}(\theta^{-1}) - 1 = k^{-1} - 1.$$

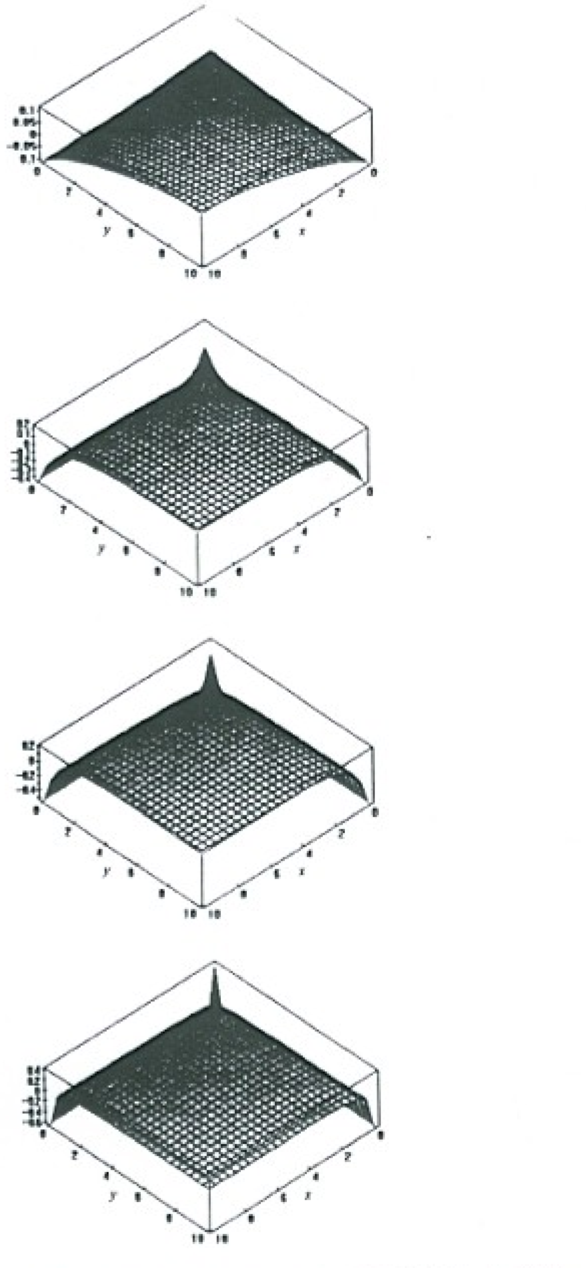

Figure 5. Graph of $H(x, y)$ for $\theta = 0.1; 1; 2; 10$ for the density (2.16).

where, as in previous case, $k = \theta \exp(-\frac{1}{\theta})/\mathrm{Ei}(\frac{1}{\theta})$, $\mathrm{Ei}(u) := \int_u^\infty v^{-1} e^{-v} dv$. Thus

$$H(x, y) = \frac{k^{-1} - 1 + \frac{\theta^2 (x^2+x-1)}{(1+\theta x)^2} \frac{(y^2+y-1)}{(1+\theta y)^2}}{\sqrt{1 + \frac{\theta^2 (x^2+x-1)}{(1+\theta x)^4}} \sqrt{1 + \frac{\theta^2 (y^2+y-1)}{(1+\theta y)^4}}}. \tag{2.18}$$

In the Table 5 some numerical values of $H(x, y)$, calculated by (2.18), are given.

Table 5.
Some numerical values of $H(x, y)$ for Gumbel's exponential distribution for $\theta = 1$ and $\theta = 2$

x	y	$H(x, y)$	x	y	$H(x, y)$	x	y	$H(x, y)$	x	y	$H(x, y)$
$\theta = 1$ ($k = 1.677$, $\rho = -0.404$, $x^* = y^* = 0.618$, $H(x^*, y^*) = -0.404$)											
0	0	0.149	1	1	−0.302	2	5	0.02	3	10	0.081
0	1	−0.308	1	5	−0.115	2	10	0.041	3	100	0.095
0	5	−0.367	1	10	−0.093	3	3	0.032	4	4	0.07
0	10	−0.36	2	2	−0.055	3	6	0.069	4	10	0.098
$\theta = 2$ ($k = 2.167$, $\rho = -0.539$, $x^* = y^* = 0.618$, $H(x^*, y^*) = -0.539$)											
0	0	0.138	0	10	−0.245	1	10	−0.328	2	10	-0.236
0	1	−0.187	1	1	−0.444	2	2	−0.281	3	3	-0.216
0	5	−0.243	1	5	−0.335	2	5	−0.243	3	10	−0.212

The values of $H(x, y)$ for this distribution are different from these presented in previous section. This is due to the fact that the marginals (but not the conditional distributions) are exponential.

3. AN ESTIMATE $H(X, Y)$

We shall briefly indicate here the manner in which $H(x, y)$ can be estimated from the data available. Evidently our proposal requires further investigation. Formula (1.3) suggests a possible estimator for local dependence function $H(x, y)$. In 1964 Nadaraya and Watson independently proposed the following estimate for the regression functions $E(X \mid Y = y)$ and $E(Y \mid X = x)$:

$$A_X^{(n)}(y) = \frac{\sum_{i=1}^{n} X_i K(\frac{y-Y_i}{h_n})}{\sum_{i=1}^{n} K(\frac{y-Y_i}{h_n})} \quad \text{and} \quad A_Y^{(n)}(x) = \frac{\sum_{i=1}^{n} Y_i K(\frac{x-X_i}{h_n})}{\sum_{i=1}^{n} K(\frac{x-X_i}{h_n})},$$

where (X_i, Y_i), $i = 1, 2, \ldots, n$ are the data, K is a kernel function, an integrable function with short tails, and $h_n \to 0$ is a width sequence tending to zero at appropriate rates. Now a "natural" estimate for Pearson correlation coefficient is given by:

$$\rho^{(n)} = \frac{n \sum X_i Y_i - \sum_i X_i \sum_j Y_j}{\sqrt{n \sum_i X_i^2 - (\sum_i X_i)^2} \sqrt{n \sum_i Y_i^2 - (\sum_i Y_i)^2}}.$$

Therefore we suggest the following estimate for $H(x, y)$:

$$H^{(n)}(x, y) = \frac{\rho^{(n)} + \frac{(\overline{X}-A_X^{(n)}(y))(\overline{Y}-A_Y^{(n)}(x))}{S_X S_Y}}{\sqrt{1+\frac{\overline{X}-A_X^{(n)}(y)}{S_X^2}}\sqrt{1+\frac{\overline{Y}-A_Y^{(n)}(x)}{S_Y^2}}}, \tag{3.1}$$

where $\overline{X} = \frac{1}{n}\sum_i X_i$, $\overline{Y} = \frac{1}{n}\sum_i Y_i$, $S_X^2 = \frac{1}{n-1}\sum_i (X_i-\overline{X})^2$ and $S_Y^2 = \frac{1}{n-1}\sum_i (Y_i-\overline{Y})^2$. Asymptotic bias and variance properties of $H^{(n)}(x, y)$ follow from the properties of estimators $\rho^{(n)}$, $A_X^{(n)}(y)$ and $A_Y^{(n)}(x)$ which are widely studied in the literature.

REFERENCES

Abrahams, J. and Thomas, J. B. (1994). A note on the characterization of bivariate densities by conditional densities. *Comm. Statist. Theory Meth.* **13**, 395–400.

Arnold, B. C. and Strauss, D. (1988). Bivariate distributions with conditionals in prescribed exponential families. *J. Roy. Statist. Soc. Ser. B* **53**, 365–375.

Bjerve, S. and Doksum, K. (1993). Correlation curves: measures of association as functions of covariate values. *Ann Statist.* **21**, 890–902.

Blyth, S. (1994). Measuring local association: an introduction to the correlation curve. *Sociol. Meth.* **24**, 171–97.

Gumbel, E. J. (1960). Bivariate exponential distributions. *J. Amer. Stat. Assoc.* **55**, 698–707.

Holland, P. W. and Wang, Y. J. (1987). Dependence function for continuous bivariate densities. *Commun. Statist. A* **16**, 863–76.

Nadaraya, E. A. (1964). On estimating regression. *Theory Probab. Appl.* **9**, 141–142.

Johnson, N. L. and Kotz, S. (1975). On some generalized Farlie–Gumbel–Morgenstern distributions. *Comm. Statist.* **4**, 415–427.

Johnson, N. L. and Kotz, S. (1977). On some generalized Farlie–Gumbel–Morgenstern distributions-II: Regression, correlation and further generalizations. *Comm. Statist.* **6**, 485–496.

Jones, M. C. (1996). The local dependence function. *Biometrika* **83**, 4, 899–904.

Jones, M. C. (1998). Constatnt local dependence. *J. Multivariate Anal.* **64**, 148–155.

Lee, M. L. T. (1996). Properties and applications of the Sarmanov family of bivariate distributions. *Comm. Statist. Theory Meth.* **25**(6), 1207–1222.

Watson, G. S. (1964). Smooth regression analysis. *Sankya Ser. A* **26**, 359–372.